

**IN THE UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF TEXAS
HOUSTON DIVISION**

SUPERSPEED, LLC,

Plaintiff,

v.

GOOGLE, INC.,

Defendant.

§
§
§
§
§
§
§
§
§

Civil Action No. 4:12-cv-01688

**JOINT STIPULATION AND [PROPOSED] ORDER REGARDING ELECTRONIC
STORED INFORMATION, DISCOVERY, AND THE FORMAT OF DOCUMENT
PRODUCTIONS**

This Joint Stipulation and [Proposed] Order Regarding Electronically Stored Information, Discovery, and the Format of Document Productions shall govern the parties in the above-captioned case (the “Litigation”).

I. GENERAL PROVISIONS

A. The parties will make reasonable efforts to prepare responsive and nonprivileged data for production in accordance with the agreed-upon specifications set forth below. These specifications apply to hard copy documents or electronically stored information (“ESI”) which are to be produced in the first instance in this litigation.

B. **SECURITY.** Both parties will make reasonable efforts to ensure that any productions made are free from viruses and provided on encrypted media for submission.

C. **CONFIDENTIALITY DESIGNATION.** Responsive documents in TIFF format will be identified with the appropriate confidentiality designations in accordance with the Protective Order in this matter. Each responsive document produced in native format will have

its confidentiality designation identified in the filename of the native file.

D. NON-STANDARD FILES. The format of production of non-standard electronic files, large oversized documents, etc. will be discussed before production to determine the optimal production format.

E. PRODUCTION MEDIA. Documents shall be produced on external hard drives, readily accessible computer(s) or electronic media ("Production Media"). Each piece of Production Media shall identify a production number corresponding to the production volume (e.g., "VOL001", "VOL002"), as well as the volume of the material in that production (e.g. "-001", "-002"). Each piece of production media shall also identify: (1) the producing party's name; (2) the production date; and (3) the Bates Number range of the materials contained on the Production Media.

II. DATA PROCESSING

A. KEYWORD SEARCHING. To the extent that keywords are used in limiting the universe of potentially responsive documents to be reviewed, the parties shall meet and confer to try to develop a mutually agreeable list of search terms and protocols prior to the production of documents. Any search methodology employed must open compound and nested files and decompression of archives. The search utilities employed must support the use of stemmed searches (e.g. using ~ to include variations on a keyword) and Boolean searches.

B. CULLING\FILTERING. Each party will use its best efforts to filter out common system files and application executable files by using a commercially reasonable hash identification process. Hash values that may be filtered out during this process are located in the National Software Reference Library ("NSRL") NIST hash set list. Additional culling of system file types based on file extension may include, but are not limited to: WINNT, LOGS, DRVS,

MP3, C++ Program File (c) , C++ Builder 6 (cpp), Channel Definition Format (cdf), Creatures Object Sources (cos), Dictionary file (dic), Executable (exe), Hypertext Cascading Style Sheet (css), JavaScript Source Code (js), Label Pro Data File (IPD), Office Data File (NICK), Office Profile Settings (ops), Outlook Rules Wizard File (rwz), Scrap Object, System File (dll), Temporary File (tmp), Windows Error Dump (dmp), Windows Media Player Skin Package (wmz), Windows NT/2000 Event View Log file (evt), Python Script files (.py, .pyc, .pud, .pyw), Program Installers.

C. The parties agree to use commercially accessible software to open password-protected or encrypted files.

D. **DEDUPLICATION.** A party is only required to produce a single copy of a responsive document and a party may de-duplicate responsive ESI (based on MD5 or SHA-1 hash values at the document level) across Custodians. For emails with attachments, the hash value is generated based on the parent/child document grouping. A party may also de-duplicate “near-duplicate” email threads as follows: In an email thread, only the final-in-time document need be produced, assuming that all previous emails in the thread are contained within the final message. Where a prior email contains an attachment, that email and attachment shall not be removed as a “near-duplicate.” To the extent that de-duplication through MD5 or SHA-1 hash values is not possible, the parties shall meet and confer to discuss any other proposed method of de-duplication.

III. PRODUCTION OF HARD COPY DOCUMENTS

A. **TIFFs.** Documents that exist in hard copy format only shall be scanned and produced as single page Group IV TIFFs, with at least 300 dots per inch (dpi). Each TIFF image shall be named according to the corresponding Bates number associated with the document. Each

image shall be branded according to the Bates number and agreed upon confidentiality designation. TIFFs shall show all text and images that would be visible to a user of the hard copy documents.

B. OCR TEXT FILES. A commercially acceptable technology for optical character recognition “OCR” shall be used for all scanned, hard copy documents. OCR text shall be provided as a single text file for each document, and the filename itself should match its respective TIFF filename. The text files will not contain the redacted portions of the documents.

C. DATABASE LOAD FILES/CROSS-REFERENCE FILES. Documents should be provided with (a) an ASCII delimited data file (.txt, .dat, or .csv), and (b) an image load file that can be loaded into commercially acceptable production software (e.g., Concordance, Summation). Each TIFF in a production must be referenced in the corresponding image load file. The total number of documents referenced in a production’s data load file should match the total number of designated document breaks in the Image Load file(s) in the production.

D. CODING FIELDS. To the extent it is readily available, the following information shall be produced in the delimited data file accompanying hard copy documents: (a) BEGBATES, (b) ENDBATES, (c) CUSTODIAN, (d) CONFIDENTIALITY, and (e) REDACTED.

E. BATES NUMBERING. All images must be assigned a unique Bates number that is sequential within a given document and across the production sets. The Bates numbering should start with a unique prefix identifying the producing party and should sequentially number the images being produced.

F. UNITIZING OF DOCUMENTS. In scanning paper documents, distinct documents should not be merged into a single record, and single documents should not be split

into multiple records (i.e., paper documents should be logically unitized). The parties will use reasonable efforts to unitize documents correctly.

G. IDENTIFICATION OF PAPER DOCUMENTS. The parties will utilize best efforts to ensure that paper records for a particular Custodian that are included in a single production are produced in consecutive Bates stamp order. The parties will identify which documents in a production are scanned paper documents either in the cover letter accompanying the production or in a coding field titled "Paper (Y/N)."

IV. PRODUCTION OF ELECTRONICALLY STORED INFORMATION

A. METADATA FIELDS AND PROCESSING. Each of the metadata and coding fields set forth in Appendix 1 that can be extracted shall be produced for that document. With the exception of prior production materials for which metadata is not available, the parties are not obligated to populate manually any of the fields in Appendix 1 if such fields cannot be extracted from a document, with the exception of the following: (a) BEGBATES, (b) ENDBATES, (c) BEGATTACH, (d) ENDATTACH; (e) CUSTODIAN, (f) CONFIDENTIALITY, and (g) REDACTED, which should be populated by the party or the party's vendor. The parties are not obligated to manually populate any of the fields in Appendix 1 for prior production materials for which metadata is not readily available. The parties will make reasonable efforts to ensure that metadata fields automatically extracted from the documents are correct, however, the parties acknowledge that such metadata may not always be accurate and might instead contain irrelevant or incorrect information generated during the collection process. Parties may request other native files be produced as described in Section IV.I. below.

B. TIFFs. Unless excepted below, single page Group IV TIFFs should be provided, at least 300 dots per inch (dpi). Each TIFF image file should be one page and named according

to the unique bates number, followed by the extension “.TIF”. Original document orientation should be maintained (i.e., portrait to portrait and landscape to landscape).

C. TEXT FILES. For each document, a text file should be provided along with the image files and metadata. The text of native files should be extracted directly from the native file. However, if a document has been redacted or does not contain extractable text, OCR of the redacted document will suffice in lieu of extracted text.

D. DATABASE LOAD FILES/CROSS-REFERENCE FILES. (a) an ASCII delimited data file (.txt, .dat, or .csv), and (b) an image load file that can be loaded into commercially acceptable production software (e.g., Concordance, Summation). Each TIFF in a production must be referenced in the corresponding image load file. The total number of documents referenced in a production’s data load file should match the total number of designated document breaks in the Image Load file(s) in the production. The total number of documents in a production should match the total number of records in the database load file. Each TIFF in a production must be referenced in the corresponding image load file. The total number of pages referenced in a production’s image load file should match the total number of TIFF files in the production.

E. BATES NUMBERING. All images must be assigned a unique Bates number that is sequential within a given document and across the production sets. The Bates numbering should start with a unique prefix identifying the producing party and should sequentially number the images being produced.

F. PRESENTATIONS. The parties shall take reasonable efforts to process presentations (MS PowerPoint, Google Presently/Punch) with hidden slides and speaker’s notes unhidden, and to show both the slide and the speaker’s notes on the TIFF image.

G. SPREADSHEETS. TIFF images of spreadsheets need not be produced unless redacted, in which instance, spreadsheets will be produced in TIFF with OCR. Native copies of spreadsheets should be produced with a link in the NativeLink field, along with extracted text and applicable metadata fields set forth in Appendix 1. A TIFF placeholder indicating the document was provided in native format should accompany the database record. If a spreadsheet has been redacted, TIFF images and OCR of the redacted document will suffice in lieu of a native file and extracted text. The parties will make reasonable efforts to ensure that any spreadsheets that are produced only as TIFF images are formatted so as to be readable.

H. PROPRIETARY FILES. To the extent a response to discovery requires production of ESI accessible only through proprietary software, the parties should continue to preserve each version of such information. The parties shall meet and confer to finalize the appropriate production format.

I. REQUEST(S) FOR ADDITIONAL NATIVE FILES. If good cause exists to request production of specified files, other than those specifically set forth above, in native format, the party may request such production and provide an explanation of the need for native file review, which request shall not unreasonably be denied. Any native files that are produced should be produced with a link in the NativeLink field, along with all extracted text and applicable metadata fields.

J. REDACTION OF INFORMATION. If documents are produced containing redacted information, the producing party shall supply a list of the documents for any such claim(s) of privilege, indicating the grounds for the redaction and the nature of the redacted material (e.g., privilege, trade secret, privacy). During the course of the litigation, an electronic copy of the originally, unredacted data shall be securely preserved in such a manner so as to

preserve without modification, alteration or addition the content of such data including any metadata therewith. If scanning is necessary to create a redacted version of a document, the producing party shall produce a corresponding load file identifying the custodian from whom the document was collected and such documents shall subject the TIFF files to Optical Character Recognition (“OCR”) process if the TIFF files do not already have extractable text, such that the processed files become searchable.

V. PROCESSING OF THIRD-PARTY DOCUMENTS

A. A party that issues a non-party subpoena (“Issuing Party”) shall include a copy of this Stipulation with the subpoena and state that the parties to the litigation have requested that third-parties produce documents in accordance with the specifications set forth herein.

B. The Issuing Party shall produce any documents obtained pursuant to a non-party subpoena to the opposing party.

C. If the non-party production is not Bates-stamped, the Issuing Party will endorse the non-party production with unique prefixes and Bates numbers prior to producing them to the opposing party.

D. Nothing in this stipulation is intended to or should be interpreted as narrowing, expanding, or otherwise affecting the rights of the parties or third-parties to object to a subpoena.

VI. SEARCHING

A. CUSTODIAL PRODUCTION. The parties shall meet and confer to develop a list of custodians and search terms with which the parties will conduct electronic searches, as appropriate, using such agreed search term lists of custodial documents. Each requesting party shall limit its request for custodial production to a total of five custodians per producing party. The parties may jointly agree to modify this limit without the Court’s leave. In the absence of

such agreement, the requesting party may petition the Court for up to five additional custodians per producing party, upon a showing of a distinct need.

Search terms will be employed for the searching of custodial data. The total search terms will be narrowly tailored to particular issues and limited to twenty terms (or as otherwise modified by agreement by the Parties). A conjunctive combination of multiple words or phrases (e.g., “computer” and “system”) narrows the search and shall count as a single search term. A disjunctive combination of multiple words or phrases (e.g., “computer” or “system”) broadens the search, and thus shall count as a separate search term unless they are variants of the same word. Should a party serve custodial production requests or search term lists beyond the limits agreed to by the Parties or granted by the Court pursuant to this paragraph, the requesting party shall bear all reasonable costs caused by such additional discovery.

B. Locations That Will Be Searched for Responsive Documents. The parties will search any electronic files or folders, or other parts of media, including any internal and external hard drives and other ESI venues (including, but not limited to, recordable optical media, media cards, thumb drives, non-volatile memory, floppy disks, work desktop and laptop computers, email servers, intranet servers, network shares, public data shares and/or database servers) for each identified Custodian that the Custodian reasonably anticipates may contain non-duplicative Responsive Documents.

C. The parties agree to search central repositories, including central databases, or relevant portions thereof to the extent that the party reasonably anticipates they contain non-duplicative Responsive Documents. The parties agree to meet and confer to limit the scope of production from central repositories if the search of central repositories (or relevant portions thereof) that the producing party anticipates contain Responsive Documents is unduly

burdensome or is likely to be unreasonably inefficient in identifying relevant documents. Specifically, the parties recognize that certain repositories, by their nature, may not effectively or reasonably be searched using electronic search strings, and the parties agree to notify each other of any such repositories that contain Responsive Documents. The parties will then meet and confer to discuss the collection of Responsive Documents from such repositories, including potentially using other effective collection methodologies.

D. Locations That Will Not Be Searched for Responsive Documents. The following locations will not be searched, and as such need not be preserved: information stored on personal digital assistants, mobile phones, voicemail systems, instant messaging systems, and automated disaster recovery backup systems and/or disaster recovery backup tapes. In addition, the parties agree that only sent and received custodial emails will be searched. Notwithstanding the foregoing, the parties agree that Responsive Documents that a Custodian indicates are stored on an archival storage medium that the Custodian can readily identify and locate, that cannot be located in any other repository of information, and that can reasonably be searched, will be searched. In addition, nothing in this paragraph shall limit a receiving party's right to request from a producing party more information about the nature of and burden associated with obtaining documents from a particular location. The parties further recognize their obligations to preserve any potentially relevant sources of data, whether live or in archival form, for purposes of this litigation.

E. SOURCE CODE. To the extent relevant to the Litigation, source code will be made available for inspection pursuant to the terms of the Protective Order. The parties agree that the search terms will not be applied to source code.

VII. MISCELLANEOUS PROVISIONS

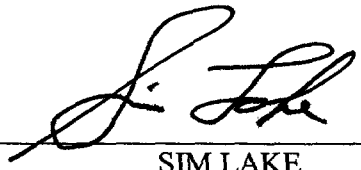
A. OBJECTIONS PRESERVED. Nothing in this protocol shall be interpreted to require disclosure of irrelevant information or relevant information protected by the attorney-client privilege, work-product doctrine, or any other applicable privilege or immunity. The parties do not waive any objections as to the production, discoverability, admissibility, or confidentiality of documents and ESI.

B. No Effect on Cost Shifting. Nothing in this Agreement shall affect, in any way, a producing party's right to seek reimbursement for costs associated with collection, review, and/or production of documents or ESI.

C. NO WAIVER. Pursuant to Federal Rule of Evidence 502(d), the inadvertent production of a privileged or work product protected ESI is not a waiver in the pending case or in any other federal or state proceeding. Moreover, the mere production of ESI in a litigation as part of a mass production shall not itself constitute a waiver for any purpose.

PURSUANT TO STIPULATION, IT IS SO ORDERED:

September 7, 2012


SIM LAKE
UNITED STATES DISTRICT COURT JUDGE

STIPULATED AND AGREED TO:

September 6, 2012

By: /s/Douglas E. Lumish [by permission]
Douglas E. Lumish (*pro hac vice*)
Cal. Bar No. 183863
dlumish@kasowitz.com
KASOWITZ BENSON TORRES &
FRIEDMAN, LLP
333 Twin Dolphin Drive, Suite 200
Redwood Shores, California 94065
Telephone: (650) 453-5170
Facsimile: (650) 453-5171

Attorney-in-Charge for Defendant
Google Inc.

Gabriel S. Gross (*pro hac vice*)
Cal. Bar. No. 254672
ggross@kasowitz.com
Joseph H. Lee (*pro hac vice*)
Cal. Bar. No. 248046
jlee@kasowitz.com
Parker C. Ankrum (*pro hac vice*)
Cal. Bar. No. 261608
pankrum@kasowitz.com
Rebecca Unruh (*pro hac vice*)
Cal. Bar. No. 267881
runruh@kasowitz.com
KASOWITZ BENSON TORRES & FRIEDMAN, LLP
333 Twin Dolphin Drive, Suite 200
Redwood Shores, California 94065
Telephone: (650) 453-5170

David J. Beck
Bar No. 00000070
Michael Richardson
Bar No. 24002838
Robert Ford
Bar No. 24074219
Beck, Redden & Secrest, LLP
One Houston Center
1221 McKinney Street, Suite 4500
Houston, Texas 77010-2010

September 6, 2012

By: /s/Neal S. Manne [by permission]

Neal S. Manne
State Bar No. 12937980
nmanne@susmangodfrey.com
SUSMAN GODFREY L.L.P.
1000 Louisiana Street, Suite 5100
Houston, Texas 77002
Telephone: (713) 651-9366
Facsimile: (713) 654-6666

Attorney-in-Charge for Plaintiff
SuperSpeed Software, LLC

Max L. Tribble Jr.
State Bar No. 20213950
mtribble@susmangodfrey.com
SUSMAN GODFREY L.L.P.
1000 Louisiana Street, Suite 5100
Houston, Texas 77002
Telephone: (713) 651-9366
Facsimile: (713) 654-6666

Kathryn P. Hoek
California Bar No. 219247
khoek@susmangodfrey.com
SUSMAN GODFREY L.L.P.
1901 Avenue of the Stars, Suite 950
Los Angeles, CA 90067
Telephone: (310) 789-3100
Facsimile: (310) 789-3150

Attorneys for Plaintiff
SuperSpeed Software, LLC

CERTIFICATE OF SERVICE

I hereby certify that on September 6, 2012, I electronically filed the foregoing document with the clerk of the court for the U.S. District Court, Southern District of Texas, using the electronic filing system of the court. The electronic case filing system sent a "Notice of Electronic Filing" to the attorneys of record who have consented in writing to accept this Notice as service of this document by electronic means.

/s/Parker Ankrum
Parker Ankrum

Appendix 1: ESI Metadata and Coding Fields

A. Image Load File shall contain the following comma-delimited fields:
 BEGBATES, VOLUME, IMAGE FILE PATH, DOCUMENT BREAK, FOLDER
 BREAK, BOX BREAK, PAGE COUNT

B. Metadata Load File shall be delimited according to the following characters:
 o Delimiter = D(ASCII:0020)
 o Text-Qualifier = ¨ (ASCII:00254)

C. The following fields will appear in the metadata load file in the order displayed below:

Field Name	Field Description
BEGBATES	Beginning Bates number as stamped on the production image
ENDBATES	Ending Bates number as stamped on the production image
BEGATTACH	First production Bates number of the first document in a family
ENDATTACH	Last production Bates number of the last document in a family
CUSTODIAN	Individual from whom the documents originated
NATIVELINK	Native File Link (Excel files only)
SUBJECT	Subject line of email
DATESENT	Date email was sent (format: MM/DD/YYYY)
TIMESENT	Time email was sent
TO	All recipients that were included on the "To" line of the email
FROM	The name and email address of the sender of the email
CC	All recipients that were included on the "CC" line of the email
BCC	All recipients that were included on the "BCC" line of the email
AUTHOR	Any value populated in the Author field of the document properties
FILENAME(Edoc only)	Filename of an electronic document
DATEMOD (Edoc only)	Date an electronic document was last modified (format: MM/DD/YYYY)
DATECREATED (Edoc only)	Date the document was created (format: MM/DD/YYYY)